AISR NNG06GE59G

On-the-fly and Grid Analysis of Astronomical Images for the Virtual Observatory

PI:
Andrew Ptak
Johns Hopkins University

Introduction

This AISR project combines two previous AISR-funded efforts, XAssist and WESIX. XAssist (http://www.xassist.org) is a program that automatically analyzes *Chandra*, *XMM-Newton*, and *ROSAT* data. XAssist is particularly adept at processing large numbers of datasets, and accordingly has been used to produce pipelines for these missions. The *Chandra* and *XMM-Newton* pipeline source lists are included as tables that can searched at the HEASARC (http://heasarc.gsfc.nasa.gov). WESIX (Web Enabled Source Identification with XMatching; http://nvogre.phyast.pitt.edu:8080/wesix) is a web service that runs the source-detection program SExtractor on an uploaded image. The resultant source list is then cross-matched with selected catalogs. The main goals of this project are to develop a set of web services for the XAssist functionality to allow X-ray analysis to be done on-the-fly and in a distributed fashion, and to expand the WESIX web service to also include X-ray images. This is a natural pairing since WESIX is an existing web service that is already part of the Virtual Observatory for non-interactive analysis of optical images, and XAssist has been developed from the beginning for non-interactive usage.

Status

XAssist

Several students were hired to work on the core code of XAssist and to improve the XAssist web site. An off-line version of the web site is being produced and will "go live" pending an upgrade of the operating system of the main web server. The XAssist code base is being refactored to improve its efficiency and to allow individual functions to be isolated. The latter step is necessary to allow the functions to be exposed as web services. This process has also facilitated more advanced user interfaces to the XAssist program (the existing interface was admittedly not very user friendly). The current version now has a command-driven interface that will in turn allow for the development of a graphical user interface. Since there will in general be a correspondence between user options and controls in a graphical interface and web services, we are exploring the possibility of having an external graphical control program that "attaches" to running instances of XAssist. This will be particularly useful in a grid environment to allow users to monitor the process of XAssist jobs running on different computers.

While restructuring the code we have been updating and, as necessary, replacing the various external packages used by XAssist. For example, the pyfits package maintained at the Space Telescope Institute is now being used for manipulating FITS files. The web template package Cheetah is now being used to produce improved web report pages (which in turn are the main mechanism for exploring the XAssist pipelines). We are testing the use of the Psycho Python compiler to improve the overall speed of XAssist and the image-fitting program ximgfit (also written primarily in Python). Another key improvement is that there is now preliminary support for the *Suzaku* mission.

WESIX

The co-I A. Connolly is supervising work on WESIX. Two key improvements this year are the addition of aperture photometry for determining magnitudes and the use of additional web services to perform automatic astrometry corrections. We have also begun to derive an updated interface to WESIX to allow for the submission of X-ray as well as optical images. The main modification will be to add parameters that are relevant to X-ray data, which of course will be ignored when only an optical image is uploaded. The WESIX project is also transitioning to a new Java web service framework named XFIRE (http://xfire.codehaus.org/). The current version uses Apache Axis, however this library had difficult in transferring FITS images and in communicating properly with other VO web services.

Current Work

There are four main bottlenecks in the processing of X-ray data by XAssist: 1) reprocessing the data to incorporate the latest calibration, 2) source detection, 3) fitting the image of each source in order to ascertain the spatial extent and to determine efficient source extraction regions for spectra, and 4) generating instrumental responses for each source to properly determine the flux, and, for sources with enough counts, to perform spectral fitting. These steps help to make results of XAssist be more reliable, but take too much time for an on-the-fly processing system. Therefore we devised a plan to allow for quicker response times.

If a requested field has already been processed, then the pipeline data can be accessed directly and these bottlenecks are obviously avoided. If not, but the data for the field requested is public or uploaded by the user, then two jobs will be spawned. The first will perform quick-look analysis of the data where reprocessing is not done (and is often only a marginal improvement), SExtractor will be used for source detection, which runs in seconds. The spatial fitting step will also be skipped and it will be assumed that sources are unresolved, i.e., model point-spread functions will be used for the source extent. Finally, pre-computed responses will be used to derive fluxes. The second job will add the field to the XAssist pipeline queue with high priority to perform full processing. XAssist is also being modified to allow only a subset of a field to be processed, which of course will shorten the processing time.

To calibrate these quick-look steps we are running simulations of point sources across the field of view (FOV) for *Chandra* and *XMM-Newton*. We are running SExtractor on images extracted from the simulation to tune its performance. The simulated data will then be run through the usual XAssist processing steps, which will produce lookup tables that will be used to quickly determine the spatial extent and counts-to-flux conversion for a source at a given position on the detector. We are also reprocessing all public *Chandra* ACIS imaging observations and *XMM-Newton* data. This will take several months to execute but will allow reprocessed data to be used when quick-look processing is needed. After this is completed the full processing will be performed as part of the XAssist pipelines. As new data become public, they will similarly be reprocessed and then added to the pipeline queue. This procedure was discussed in a poster given at the October 2006 ADASS meeting, and a flow chart is shown in Figure 1.

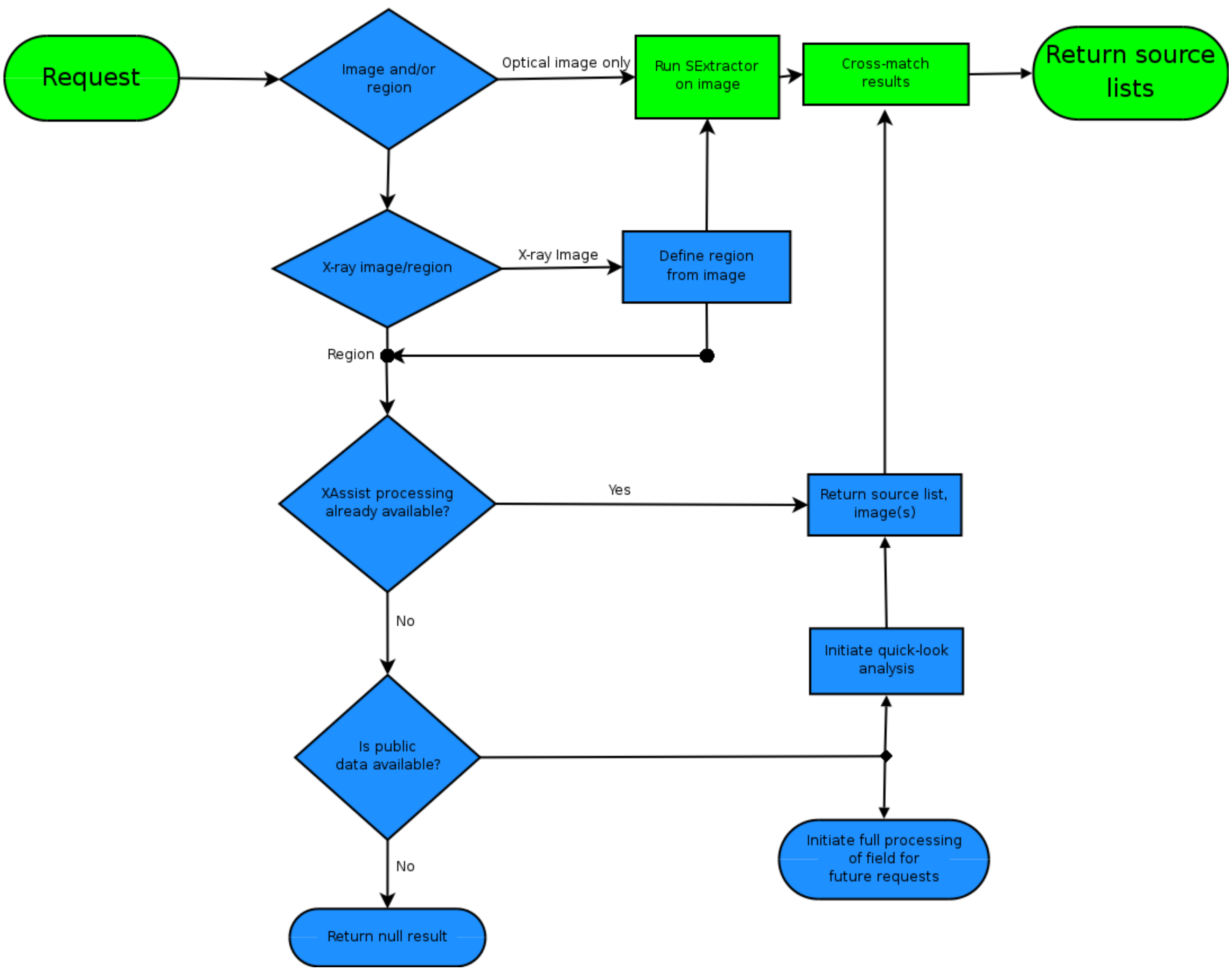

Figure 1: Flowchart for the combined XAssist/WESIX web service.

Future Work

We anticipate the calibration work to be completed by the beginning of the summer of 2007. Since the reprocessing functionality of XAssist includes generating exposure maps, we will be able to quickly implement a fingerprint service to determine if a given source position is in the FOV of *Chandra* and *XMM-Newton* observations. We anticipate that this will be online by the end of the summer of 2007. We also anticipate the development of web services for the basic functionality of XAssist by the end of 2007, namely web services to request the full processing of a dataset, to retrieve the properties of sources within a given region of a processed field (if any have been detected), to compute upper-limits at a given position, and to perform quick-look analysis of a given dataset (to be directly incorporated into the WESIX web service). Over this time period we will continue to improve the performance of both XAssist and WESIX separately. Both the quick-look and full processing functionality will be implemented for *Suzaku* data, and, time permitting, also *Swift* XRT data.